1-17-2008

Set-Switching and Learning Transfer

C. Dustin Johnson

SET-SWITCHING AND LEARNING TRANSFER

by

C. DUSTIN JOHNSON

Under the Direction of David A. Washburn

ABSTRACT

In this experiment I investigated the relationship between set-switching and transfer learning, both of which presumably invoke executive functioning (EF), which may in turn be correlated with intelligence. Set-switching was measured by a computerized version of the Wisconsin Card Sort Task. Another computer task was written to measure learning-transfer ability. The data indicate little correlation between the ability to transfer learning and the capacity for set-switching. That is, these abilities may draw from independent cognitive mechanisms. The major difference may be requirement to utilize previous learning in a new way in the learning-transfer task.

INDEX WORDS: Learning Transfer, Transfer Index, Wisconsin Card Sort Task, Set-switching, Intelligence, Executive Functioning

SET-SWITCHING AND LEARNING TRANSFER

by

C. DUSTIN JOHNSON

An Honors Thesis Submitted in Partial Fulfillment of the

Requirements for Graduation with Undergraduate Research Honors

in the College of Arts and Sciences

Georgia State University

2007

SET-SWITCHING AND LEARNING TRANSFER

by

C. DUSTIN JOHNSON

Honors Thesis Director: David A. Washburn
Honors Program Director: Robert D. Sattelmeyer

Electronic Version Approved:

Honors Program
College of Arts and Sciences
Georgia State University
May, 2008

Acknowledgments

I would like to thank David A. Washburn, Ph. D. for his patience, persistence, and continued guidance in this project. Thanks also to Natasha Schultz for her help with the execution of the experiment, data analysis, and thesis writing. I give special thanks to the National Institute of Child Health and Human Development and the U.S. Army Medical Corp and Materiel Command for their grants that partially funded this project. Lastly, thanks to Georgia State University and the Department of Psychology for providing the facilities for my work.

Table of Contents

Set-switching and Learning Transfer

There exists a long-standing debate as to the precise nature of intelligence and how it should be measured. Gardner's (1983) theory of multiple intelligences suggests that there are seven distinct types of intellectual abilities. Sternberg (1985) postulated three sets of components constituting cognitive capacity. Perhaps the most widely accepted model is Spearman's (1904) general intelligence model, which is hierarchical and includes several broad abilities, such as visuo-spatial, memory, and verbal ability (Kane & Brand, 2006). At the top of the hierarchy is a general intelligence factor, *g*. This factor has been used successfully in accounting for performance on a wide variety of cognitive tests, although much mystery still remains as to the nature of underlying general mechanisms (Kolata, Light, Grossman, Hale, & Martzel, 2007).

The general intelligence model shares common features with Baddeley and Hitch's (1974) model of working memory, which includes specialized storage systems (e.g., a phonological loop, a visuo-spatial sketchpad) coordinated by a general central executive. Research findings support a strong relationship between *g* and working memory. Miyake and Friedman (2001) found that executive functioning combined with visuo-spatial or language skills are good indicators of *g*. Likewise, Colom and Rebollo (2005, p.624) contended that working memory and *g* are "almost isomorphic concepts." Kolata et al. (2007) found that general learning abilities were correlated to the selective attention component but not the short-term memory component of working memory. In other words, these theorists claim that working memory is closely related to, and perhaps even synonymous with, the most popular notion of intelligence, *g*. However, debate remains as to which component(s) of working memory best represent *g*.

The phonological loop is the component of working memory that stores verbal material and allows for its rehearsal and/or maintenance (Baddeley & Hitch, 1974). The visuo-spatial sketchpad is a storage facility for visual and spatial material. The central executive is responsible for actively coordinating the other two systems. This includes decision making, attention focusing, and reasoning (Logie, 1995). Because the central executive's role in working memory entails the greatest responsibility, it would seem to be a good candidate for being the most closely related to *g*.

The operation of the central executive is referred to as executive functioning (EF) and is often invoked when nonlearned responses are required (Towse & Houston-Price, 2001). The central executive can also direct attention to an important stimulus, a process called 'selective attention' (Miller & Cohen, 2001). Through this action, new learning ensues. In some cases, this new learning must necessarily overwrite old learning. An example of this is found in the Wisconsin Card Sort Task (WCST), which can be used as a measure of EF. The WCST requires participants to sort cards on the basis of some rule, and then to switch to a new rule when the task changes. The rules constantly change, creating the dual challenge of learning new rules and inhibiting the old ones. For each rule in the WCST, participants must develop a unique rule-like strategy used for responding to the stimuli, known as a response set. As rules change on subsequent problems, participants must generate and employ appropriate response sets. Rule changing (as required by tasks like the WCST) entails an ability to switch response sets, whereby one must inhibit previously learned rules and apply new ones. This cognitive operation is known as set-switching. In the WCST, for example, one may be sorting cards on the basis of color (a "color response set") and then switch to a different response set (e.g., shape).

As the coordinator of the memory systems of working memory, the central executive presumably controls set-switching. Because the studies discussed earlier indicate a close relationship between working memory and intelligence, it is possible that set-switching ability, as a potentially large component of EF, is itself a major component of *g*. Supporting this theory, Dempstser (1991) believed that inhibitory processes, like those required by set-switching, are a substantial dimension of intelligence. He supports this claim by citing neurological evidence and the requirement for suppression of nonrelevant information on many tasks. Other studies have shown a significant correlation between executive abilities and performance on general intelligence tests (Aristide, Bernardo, Grazia, & Gaspare, 2006). These studies, taken together, indicate the possibility of a close relationship between intelligence and EF as reflected in set-switching ability. However, the nature of this relationship remains unclear.

If the WCST, as a measure of EF, does predict intelligence, then a closer examination of the cognitive mechanisms involved in performing the WCST may reveal the cognitive mechanisms that influence intelligence. Performance on the WCST is measured by perseveration errors, which occur when a participant responds according to an old rule even though he or she is aware that a new rule is in effect. Thus, a perseveration error occurs when an incorrect response set is used and is an error in set-switching. Presumably, perseveration errors result from failure to inhibit the previously learned response set, failure to implement the new response set, or a combination of both. Because the WCST is thought to measure EF, determining the cause of perseveration errors would help determine how EF is related to intelligence.

These types of errors may also occur when organisms are required to adapt previous learning to solve novel problems. This capability, known as the ability to transfer learning, requires one to transform previous learning to meet new demands. The difference between the abilities of set-switching and learning transfer is subtle yet potentially significant. To set-switch, one must abstain from using an old strategy and instead implement a new one. To transfer learning, one must apply old learning (knowledge acquired through previous experience) in a new way or to a new problem. The similarities between set-switching and learning transfer are obvious: Both cases require new learning to occur through a change in response set. Because they both entail a change in response to a stimulus, they also must both involve inhibition of the previously learned responses. For set-switching, usage of the previous rule must be sufficiently inhibited so that the new response set is employed. For transfer of learning, the inhibition must be less complete because the previously learned response is the mandatory starting point for rule transformation and subsequent new response. Because set-switching and learning transfer both entail inhibition and learning, it is reasonable to speculate that the significant difference between the two may lie in the capacity to modify previous learning, found only in learning transfer. This ability implies neural plasticity and the ability to deviate from rigorous stimulus-response habit patterns.

Because of the similar demands both capacities place on EF, one would expect the ability to set-switch to correlate highly with the ability to transfer learning. Due to EF's relationship to general intelligence (as previously discussed), it is likely that set-switching and learning transfer abilities would be components of general intelligence. If this were the case, then it would be important to determine which cognitive mechanisms are responsible for the correlation between EF (as reflected in set switching) and transfer learning. In what way, if any, is the ability to inhibit previous response patterns in set-switching related to the general ability to learn and to transfer learning to new contexts? In this experiment I aimed to begin the process of understanding these underlying mechanisms. This was done by looking specifically at set-switching and learning reversal abilities, which were measured by two tasks that seemed intuitively to make very similar cognitive demands. The demands were so similar that one would expect performance on the learning transfer task to predict performance on the set-switching task. If this is the case, then transfer learning ability may be further pursued as an indicator of EF and intelligence. If little correlation is found, then these constructs may point to distinct, independent cognitive abilities.

As a conventional measure of set-switching and EF, the WCST was utilized. As a measure of learning transfer ability, an assessment has been developed and tested: the transfer index, henceforth referred to as TI. The concept of 'TI' was created by Rumbaugh (1969, Rumbaugh & Pate, 1984) to quantify intelligence, particularly of nonhuman primates, through learning reversal ability. Reversing learning is a special instance of learning transfer, where the desired response, once learned, becomes the undesired response. Hence, in order to generate the newly desired response, one must recall the previously desired response and perform in the opposite way. Thus, the desired response is a modification of previously learned behavior.

Traditional intelligence tests are subject to numerous criticisms. For instance, one never knows how much a person has learned prior to testing (e.g., as the result of culture, schooling, or previous exposure to tests) that might inflate the estimate of that person's ability to learn. In learning transfer tests, these biases are minimal because novel problems are presented that have never been learned. Every person or animal being tested begins with the same level of knowledge about the problem. As such, tests designed around the transfer index paradigm would be both unbiased and unresponsive to a person's previous educational experiences. Establishing the connection between learning transfer ability and intelligence would facilitate the design of these tests.

Rumbaugh's TI is designed to control for individual and species differences in the amount of prior education or training, familiarity with testing, and other confounds unrelated to intelligence. The procedure begins by allowing the organism to learn a novel problem to a specified performance level (Rumbaugh 1969, Rumbaugh & Pate, 1984). Once the learning criterion is reached, the problem is reversed so that the same response rule applies, but in the opposite way. For instance, if the initial rule was to respond on the basis of size by selecting the larger of two items, after reversal one is correct for selecting on the basis of size by indicating the smaller item. Rumbaugh and colleagues have shown that relatively intelligent animals (i.e., those with relatively large and encephalized brains, and that learn well in a variety of other situations) tend to apply the original rule they learned to the new conditions well, whereas less intelligent organisms tend not to reverse the response cues and rather to respond in whatever way had previously been reinforced. Moreover, previous studies by Rumbaugh indicate that the better the rule is learned prior to reversal (i.e., the higher the prereversal criterion level is set), the bigger the difference in performance between individuals with relatively good versus poor learning ability. In other words, as the requirement for prereversal accuracy is increased, post-reversal performance becomes more difficult for less intelligent organisms and relatively easier for more intelligent organisms. Presumably, this difference results from the differing levels of sophistication of rules generated by animals of varying intelligence. For animals that are more controlled by stimulus-response-associative forms of learning, the more reinforcement they have received for a certain response, the harder it is for them to abandon that response in favor of another stimulus. Such animals might generate rules that concern only the specific stimulus-response connections that have been previously experienced. With repetition, these connections strengthen in the form of habit. For animals that are capable of relational forms of learning, reinforcement strengthens the rule, not the habit to responding to any particular stimulus indicated by the rule. These animals are more likely to base their responses on rules they have abstracted from a variety of stimulus-response relationships they have previously experienced. Thus, these animals transfer learning well whenever the stimulus conditions change but the rule remains constant.

Rumbaugh (1969) quantified the TI by relating participant performance from one task or problem (unique set of stimuli) to the next with the rules changing only between problems. Specifically, the TI was the performance after the rule change divided by performance before the rule change. A higher TI implied a greater ability to adapt the old rule to the new situation.

$$TI = \frac{\text{Performance after rule change}}{\text{Performance before rule change}}$$

As a provisional measure, I employed in this experiment a modified TI that only reflects participant performance above and beyond what would have been expected for random guessing. The TI task implemented here required binary responses only; so random guessing should have yielded 50% correct. This measure was named the transfer index corrected, or TIC. As with TI, greater TIC indicated greater learning transfer. A negative TIC indicated that the participant’s performance after the rule change was worse than what would have been expected from random guessing.

$$TIC = \frac{\text{(Performance after rule change)} - \text{(performance from random guessing)}}{\text{(Performance before rule change)} - \text{(performance from random guessing)}}$$

ther researchers have found that there may exist a significant relationship between the ability to reverse learning and intelligence. For instance, Rajalakshmi and Jeeves (1965) found that reversal learning correlates more highly than discrimination learning with intelligence in children, as measured by the Wechsler Intelligence Scale for Children (WISC). This experiment sought to continue this line of inquiry. If the ability to transfer learning is discovered to be a viable measure of EF and, by extension, intelligence, then intelligence tests could be constructed utilizing TI concepts to achieve greater validity. Whereas many conventional tests often measure intelligence by how much a person knows, an intelligence test based on learning transfer ability could measure intelligence by how easily a person can transfer her learning.

The hypothesis of this experiment is that the ability to transfer learning, as measured by a Learning Reversal Task will predict the ability to switch response sets efficiently, as measured by the WCST.

Method

Participants

Forty-five Georgia State University undergraduate students participated in this study (27 were female and 18 were male). In return, they were each given one hour of credit towards fulfilling their research requirements for an introductory psychology course. Recruitment, registration for time slots, and credit dispensation were done using Sona Systems, a web-based software solution for subject pool management. All participants had normal or corrected-to-normal vision. The participants were between 17 and 45 years old and had a mean age of 20.5 years.

Apparatus and Tasks

The experiment was administered on four Dell personal computers running Windows XP with 17-inch monitors. Each computer contained two relevant Visual Basic programs: the *Learning Reversal Task* and the *WCST*, both written by David Washburn, Ph.D. of Georgia State University. The participants utilized the keyboard during the program's execution.

The *Learning Reversal Task* was a game written specifically for this experiment in which the participants had to continually adapt to its changing rules. Through the process of trial and error, the participant was expected to learn the rule and use it to make accurate determinations of which of two stimuli was correct. If the participant demonstrated sufficient prowess using the rule, then the rule was changed. At that point, the participants had to discover and learn a new rule, while mentally suppressing or altering any learning that remained of the old rule.

As shown in Figure 1, the game presented two binary strings of numbers, one of which was correct by the rule and the other incorrect. The rule was determined randomly by the computer, which generated a prototype stimulus at the beginning of each game. The prototype consisted of one or two randomly generated binary stimuli (‘1’ or ‘0’), each placed in randomly generated positions within the sequence of digits. Each sequence contained a total of four digits, including the prototype digit(s). Once the prototype values and positions were set, they did not vary until the next problem. The positions that were not fixed by the prototype were allowed to vary randomly. For example, suppose that the computer generated a prototype where the first position was a ‘1’ and the third position was a ‘0’. In this example, the second and fourth positions would vary randomly. Hence, possible stimuli included ‘**1**0**0**1’, ‘**1**1**0**1’, ‘**1**0**0**0’, and ‘**1**1**0**0’. (For illustrative purposes, the prototype digits are indicated by bold type; they were not obviously distinguished from other digits in the actual stimulus presentations.)

The prototype and its supplemental random digits were randomly presented on either the left or right side of the screen. The participant guessed which answer was correct by pressing the arrow key that corresponded with the left or right stimulus. When the participants selected the string that was in accordance with the rule, the word "CORRECT" appeared on the screen. Otherwise, their choice was wrong and "INCORRECT" displayed. After a second and a half, the feedback word disappeared and the process began again as two new stimuli appeared on the screen.

The rule changed when the participant was determined to be sufficiently proficient in her use of the current rule. The participant's acquisition of the rule was determined statistically by the percentage of correct selections over previous trials since the presentation of the new problem. For a reversal to occur, a minimum of 10 trials must have elapsed since the last reversal. In the case that 22 or more trials had elapsed since the previous reversal, only the 21 most current responses were included in the calculation. In the event that sixty trials passed without the participant acquiring the rule to the criterion level, the problem was ended and the next one was initiated. For the *Learning Reversal Task*, participants were tested to two different criterion levels. On the first ten problems, the criterion level was 67% correct binary selections for rule acquisition. On the second ten problems, they were required to score 80% correct before the rule was considered acquired.

When the criterion level was attained, be it 67% or 80% depending on the portion of the experiment, the rule was then changed in order to indicate that a new rule is in effect. The change of rule was indicated by a valence reversal of the sorting criteria. Hence, if a '1' was required at a particular spot before the rule change, then a '0' would be displayed at that spot after the rule change. If the participant continued to choose by the previous rule, then he was informed that his selection was incorrect. Here, it was assumed that the participant would conclude that the rule had changed and that he would begin a new process of trial and error in order to determine the new rule. After the reversal trial, the participant was then given 10 trials with the new rule, for a total of 11 trials with the new rule. Performance on the final 10 trials, not including the rule reversal trial, was compared with pre-reversal accuracy (66.7% or 80%). The ratio of performance after the reversal to the performance before the reversal was the TI.

For the WCST, sorting rules included color (green, blue, red), number (1, 2, 3), and shape (circle, triangle, square). For each trial, the participant was shown a stimulus on the screen of random color, number, and shape. Figure 2 gives an idea of what the WCST displays on the screen. For clarity in black-and-white printing, horizontal, diagonal, and vertical bars have been substituted for green, blue, and red colors respectively. As shown in Figure 2, three green squares might have been displayed. The participant would guess the sorting rule and make a selection based on that hypothesis. If the rule was guessed to be ‘color,’ then the key ‘1,’ the selection corresponding to green would be pressed. If the rule was guessed to be ‘shape,’ then the key ‘2,’ the selection corresponding to squares would be pressed. If the rule was guessed to be ‘quantity,’ then the key ‘3,’ the key corresponding to three would be pressed. The computer then displayed ‘CORRECT’ or ‘WRONG’ based on the participant’s answer and the current rule. Then, a new stimulus was presented and the participant guessed again. By the process of elimination, the participant could determine what rule the computer was utilizing.

The rule was changed when the participant answered correctly for eight consecutive trials. Upon realizing that the old one failed to yield correct results, the participant presumably would then search for the new rule. Continuing to sort by the old rule constituted perseveration errors.

Design

Learning Reversal Task. The independent variable was the criterion level to which participants had to learn the sorting rules. In some cases, they had to achieve 67% correct responses in the previous prereversal trials (the 10 to 21 most recent trials since new stimuli were presented) in order to proceed to the next part of the experiment. In other cases, they had to achieve 80% success to move forward to the reversal trials.

The dependent variables included the number of trials to criterion (NTC) and TI ratio. NTC represented the number of trials it took participants to achieve a certain percentage of correct responses. Each participant was required to achieve two success levels, 67% and 80%. As was discussed above, TI was calculated by percentage of correct answers on the ten trials following the reversal trial, divided by the criterion level, 67% or 80%. This number indicated the relative performance of the participant after the reversal versus the performance before the reversal.

Wisconsin Card Sort Task. The dependent variables on the WCST were the number of perseveration errors (NPE) and the total number of trials. NPE represented the number of times during the post-reversal trials that a participant responded in a way that indicated judgment using the pre-reversal rule. The post-reversal period for one trial was the same as the pre-reversal period for the next trial. Reversals occurred after eight consecutive correct selections. From NPE, the mean number of perseverations per problem was calculated. The total number of trials was the number of trials required to complete all ten problems.

Procedure

Upon arrival, the participants were led to their personal computers. Some participants would perform the task in a room alone, whereas others shared a testing room with one other participant due to space constraints. Once in the room and seated at their computer, the participants were asked to sign a consent form and given an optional demographics form to complete. Then, the participants were instructed on how to initiate their tasks. For example, they were shown the icons on the computer desktop on which they should click in order to initiate their tasks. Next, participants were given a set of standardized instructions for how to perform the tasks. After verifying that the participants had no questions, the experimenter left them to execute their tasks.

First, they ran the Learning Reversal Task, then they ran the Wisconsin Card Sort Task. They were given one hour in which to complete both tasks. If they had not finished one hour, then the experimenter stopped the experiment. Before the participants left, the experimenter debriefed them, offered to answer any of their questions, and thanked them for their time.

Results

Forty-five people participated (mean age 20.5 years old, 27 females and 18 males). For the Learning Reversal Task, three participants did not learn any problems to the 67% criterion level. Four other participants did not learn any problems to the 80% criterion level. Six participants did not complete the WCST before the hour allotted to the experiment had elapsed. In total, 33 of the participants reached criterion for both the 67% and 80% learning levels for the Learning Reversal Task and completed the WCST. These results reflect only the data from these 33 participants.

Learning Reversal Task

The TI ratio for problems learned to the 67% criterion level had a mean of 0.75 and ranged from 0.50 to 1.06, $SD = 0.16$. TIC for 67% had a mean of 0.0063 and ranged from -0.97 to 1.24, $SD = 0.62$. To obtain the 67% criterion level, participants required a mean of 24 trials and ranged between 11 and 37 trials, *SD* = 6.9. In 74% of the problems, participants did not learn to criterion level. Of the 10 problems presented, participants learned a mean of 3.6 problems to the criterion level, with a range between 1 and 7 problems, *SD* = 1.5. For the 67% learning level, the median averaged TI was 0.729; the 25% and 75% percentiles were 0.582 and 0.861 respectively.

For problems learned to the 80% criterion, the TI had a mean of 0.68 and ranged from 0.37 to 0.96, *SD* = 0.12. TIC in this case had a mean of 0.16 and ranged from -0.68 to 0.89, *SD* = 0.33. Participants required a mean of 19 trials and ranged between 11 and 30 trials to reach the 80% criterion level, *SD* = 5.3. They learned a mean of 4.9 problems to the 80 % criterion level and ranged between 1 and 8 problems, *SD* = 2.3. For the 80% learning level, the median averaged TI was 0.694; the 25% and 75% percentiles were 0.597 and 0.750 respectively.

Wisconsin Card Sort Task (WCST)

For the WCST, performance also varied widely. Overall, participants answered a mean of 72% of the problems correctly and ranged between 45% and 87%, *SD* = 11%. For 10 reversals, they required a mean of 218 trials with a range between 134 and 588 trials, *SD* = 117. Three participants did not complete 10 reversals, though they continued guessing for over 1000 trials. Their data was not included in these results.

The mean total number of perseveration errors over 10 reversals for the WCST was 15 perseverations and varied between 3 and 36 perseverations, *SD* = 7.5. The median number of perseveration errors was 13.0; the 25% and 75% percentiles were 9.5 and 19.5 perseverations respectively. The mean perseverations per problem was 1.6 perseverations and varied between 0.33 and 4.0 perseverations, *SD* = 0.84. The median number of perseverations per problem was

1.44 perseverations; the 25% and 75% percentiles were 1.06 and 2.17 perseverations respectively.

Correlation between Learning Reversal Task and WCST

Multiple variables from the Learning Reversal Task and the WCST were compared. For the Learning Reversal Task, the average number required to reach criterion level and the average transfer indexes were compared to the WCST mean number of perseverations per problem and the total trials required for completion.

A regression analysis was employed to relate pre-reversal performance on the learning reversal task to the mean number of perseveration errors on the WCST. Pre-reversal performance was indicated by the mean number of trials required to reach criterion. In neither the 67% nor the 80% learning scenario, was initial learning ability of the rule significantly correlated to WCST perseveration errors. As shown in Figure 3, when the rule was learned to 67%, the correlation approached significance $F(1, 31) = 3.59$, $p = 0.068$, $R^2 = 0.104$. Results were less correlated when the rule was learned to 80%, $F(1, 31) < 0.00$, $p = .993$, $R^2 = 0.00$.

Post-reversal performance on the Learning Reversal Task (indicated by average TI was then compared by regression analysis to the mean number of perseveration errors on the WCST. For problems learned to neither the 67% nor the 80% criterion level was the TI significantly correlated to WCST perseveration errors. Thus, as shown in Figure 4, post-reversal performance in the 67% case did not predict performance on the WCST, $F(1, 31) = 0.198$, $p = 0.659$, $R^2 = 0.006$. Likewise, learning the rule to 80% also yielded insignificant results, $F(1, 31) = 0.159$, $p = .693$, $R^2 = 0.005$.

In a similar manner, TIC was compared to the mean number of perseveration errors on the WCST. Post-reversal performance in the 67% case did not predict performance on the

WCST, $F(1, 31) = 0.165$, $p = 0.688$, $R^2 = 0.005$. Likewise, learning the rule to 80% also yielded insignificant results, $F(1, 31) = 0.123$, $p = .729$, $R^2 = 0.004$.

Post-reversal performance on the Learning Reversal Task (indicated by average TI was then compared by regression analysis to the total number of trials required to complete the WCST. For problems learned to the 67% criterion level, the TI significantly predicted WCST total trials, WCST, $F(1, 31) = 10.6$, $p = 0.003$, $R^2 = 0.504$. However, learning the rule to 80% criterion did not predict the total number of WCST trials, $F(1, 31) = 2.44$, $p = .129$, $R^2 = 0.270$.

Next, an extreme-groups analysis of variance was performed to see whether grouping participants by TI performance into quartile and median splits revealed differences between the groups in the mean number of perseverations per problem on the WCST. The best and worst performers of TI (grouped by quartiles) did not differ in WCST perseverations, $F(1, 13) = 0.001$, $p = 0.976$ for the 67% criterion; $F(1, 14) = 0.157$, $p = 0.698$ for the 80% criterion. Likewise, median splitting of performers at the 67% criterion level did not reveal group differences in perseverations perseverations, $F(1, 31) = 0.922$, $p = 0.345$ for the 67% criterion; $F(1, 31) = 0.532$, $p = 0.471$ for the 80% criterion.

Although TI-based groups did not differ with respect to number of perseverations on the WCST, they did differ significantly on the total number of trials required by participants to complete the WCST. On the TI at 67% criterion, participants required significantly more trials on the WCST for bottom quartile versus the top quartile, $F(1, 13) = 4.66$, $p = 0.05$. Bottom quartile performers were ones with TI < 0.582 and required a mean of 329 trials to complete the WCST. Top quartile performers had a TI > 0.861 and required a mean 176 WCST trials. Likewise, quartile splits of performers on the TI at 80% learning revealed significant differences in total number of trials on the WCST, $F(1, 14) = 4.84$, $p = .045$. The top 25% of TI performers

(TI > 0.750) required a mean of 173 trials to complete the WCST, whereas the bottom 25% (TI < 0.597) required a mean of 314 trials.

High and low performers on the WCST, as measured by mean number of perseverations per problem, were also divided by quartile and median splits. Univariate analysis showed that top and bottom quartile groups on the WCST did not differ with respect to TI on problems at the 67% criterion level, $F(1, 14) = 0.229$, $p = 0.640$ or the 80% criterion level, $F(1, 14) = 0.171$, $p = 0.685$. Similarly, univariate analysis showed that top and bottom groups from median splits of the WCST perseveration distributions did not yield different TI scores at the 67% criterion level, $F(1, 31) = 0.373$, $p = 0.546$, or the 80% criterion level, $F(1, 31) = 1.205$, $p = 0.281$.

Discussion

Despite the similarities between the tasks and the reasonable expectation that both would related to general intelligence, set-switching was unassociated to learning transfer in this study for both the 67% nor the 80% criterion levels. It is clear from these results that either the ability to reverse learning and set-switching are independent capacities or that the tests used here were not valid measures of these abilities. Assuming the former to be the case, it is likely that set-switching and transfer learning draw from independent cognitive mechanisms. As noted in the introduction, set-switching requires response set changing whereas learning transfer involves response set modification. Although set-switching draws heavily from inhibitory processes, learning transfer also requires some level of inhibition. Possibly, the requirement of learning transfer to modify previous learning draws from a unique cognitive ability and accounts for the significant difference between the two is that abilities. It is possible that individual differences in this ability to modify learning accounted for the lack of correlation between set-switching and learning transfer. Such a conclusion is commensurate with Rumbaugh’s findings that less encephalized organisms have a reduced ability to alter previously learned responses (Rumbaugh & Pate, 1984). However, the question remains as to whether or not the individual differences in learning reversal ability in this study predict intelligence. An investigation into this relationship is the next logical step. Ideally, multiple measures of intelligence would be compared to the ability to reverse learning.

Of all the analysis done examining the relationship between the Learning Reversal Task and the WCST, only one set of variables showed significant correlation: Quartile splitting of the top and bottom 25% of performers of TI predicted the total number of trials on the WCST for both the 67% and 80% criterion cases. However, the factor typically regarded as significant for the WCST is the number of perseveration errors (a measure of the efficiency of set switching), not the number of trials required for completion (which is probably a measure of learning itself). Thus, it is unsurprising that the participants that participants who learned and transferred learning well on one task also learned faster on the WCST compared to the worst learning-transfer performers. However, because there was no significant correlation between TI and WCST perseveration errors, it is likely that WCST perseveration errors and WCST total trials indicate different things. This could be caused by instances of giving wrong responses and still not sorting by the previous rule. This is possible because each trial presents the participant with three choices, one correct for the current rule, one correct for the previous rule, and one not correct for either. Hence, it is possible to make errors and increase the total number of trials while at the same time not making perseveration errors (at least not perseveration errors concerning the *immediately* preceding problem). It is not clear, however, whether or not these nonperseveration errors are indicative of EF capacity.

One other analysis showed that there may exist a significant relationship between perseveration errors on the WCST and pre-reversal performance on the Learning Reversal Task, as indicated by the mean number of trials needed to reach the 67% criterion level. In other words, what presumably is not testing learning reversal ability (pre-reversal performance), is somewhat predicting EF as measured by WCST perseverations. This could be explained by the heavy load placed on EF by hypothesis testing in the Learning Reversal Task. If fact, the central executive could be busier during hypothesis testing, when one must keep track of old hypotheses and generate new ones, than in post-reversal trials, when it only has to recall the correct hypothesis and give the opposite response. Fewer trials required to reach criterion implies more effective hypothesis testing, which may indicate greater EF. If this is true, then prereversal performance would be a measure of EF. Because WCST perseverations supposedly measure EF as well, it makes sense that the two would be correlated.

Another possible confound could result from differences in the two computer tasks. Perhaps they invoked different aspects of working memory due to the different types of stimuli they presented. Because one task presented shape and color (subject to visual memory) whereas the other task only presented binary variables (subject to aticulatory rehearsal), different proportions of the visuo-spatial sketchpad and phonological loop may have been utilized. Future work could investigate this via a process of distraction utilizing articulatory suppression or random number generation. The presenting stimuli of the tasks could also be changed.

In this sample that there existed a wide variety of performance levels. Whereas some participants finished the tasks with near-perfect scores, others never responded in ways that indicated that they understood the rules of the tasks. Such a variation in performance could be the result differences in intelligence, familiarity with computer games, motivation, or confidence in testing scenarios. It is also noteworthy that most of the participants who did not finish a task, only failed to finish one task, not both. Only a couple of participants did not complete either task. This would be most easily explained by the different cognitive demands invoked by the two tasks as related to differences in participant strengths and weaknesses. Further research would be required to confirm this.

It was interesting that 12 of the 45 participants did not complete the Learning Reversal Task and/or the WCST. Their data were not entered into the analyses. For instance, some participants did not reach criterion on the Learning Reversal Task. Because no criterion was reached, reversal and post-reversal trials did not occur. Instead, the participants kept guessing in the prereversal stage for the duration of the experiment, generating large numbers of trials as a result. Likewise, several people never concluded the WCST. Instead, they attempted more than one thousand problems before exceeding the allotted 1-hour time period. The participants, as college students, presumably possessed the required intelligence to successfully complete both tasks. Therefore, their failures were probably due to either inadequate instruction or motivation. If insufficient instruction was the culprit, then it could be better elucidated for future participants. Also, practice rounds could be provided with automated feedback describing to participants why given responses are correct or incorrect. The practice rounds could be the same or different from the actual tasks. Another reason for culling the data of these participants was the possibility that they were not motivated to understand the rules of the game. Perhaps they desired to finish the experiment early and proceeded to answer quickly and carelessly. To increase participant motivation, the tasks could be designed so that high performers complete the tasks in a substantially shorter period of time. Participants provided with this information would be more motivated to perform well in order to complete the task early.

For the Learning Reversal Task, the data showed that participants exhibited better prereversal performance at the 80% criterion level than at the 67% criterion level in regards to the number of problems learned to the criterion level (4.9 versus 3.6 problems) and the mean number of trials to reach the criterion level (19 versus 24 trials). These results are surprising because one would expect fewer trials to be required to obtain 67% success than 80% success. Likewise, it would be expected that a greater number of problems would be learned successfully in the lower criterion case.

Another unexpected result was that participants had a better mean TI at 67% criterion level (0.75) than at 80% (0.68). In other words, their post-reversal-to-preversal performance ratio was better for problems learned to a lower success rate. At first glance, this seems contrary to Rumbaugh’s findings that more intelligent organisms (like the human participants in this study) have a better TI for higher criterion levels. However, Rumbaugh’s findings only state that this holds true relative to less intelligent organisms, between species rather than within species. Nevertheless, the present data are not consistent with the principle that more learning should lead to better transfer for intelligent organisms. If this finding were to replicate, it would undermine the logic of the TI paradigm as a measure of intelligence.

It may also the case that higher TI averages for the lower initial criterion can be an artifact of the way that TI is calculated. TI might be more intelligible if its components reflected only performance that was above and beyond random guessing. As was discussed in the introduction, the TIC measure I have proposed corrects for this effect by looking only at performance greater than what would be expected with random guessing. In this formula, a TIC = 0 would indicate equal pre- and post-reversal performance; and a perfect post-reversal performance would yield TIC = 2.9 for 67% criterion and TIC = 1.7 for 80% criterion. Thus, perfect performance in the more severe case of under-learning (i.e. 67% vs. 80% criterion) is better rewarded in the TIC score. In the Learning Reversal Task, the mean TIC is greater for the 80% criterion (0.16, which is about 9% of the optimal score) than for the 67% criterion (0.0063, or less than 1% of the optimal score). As readily evident by these numbers, both of the effects are modest and learning was better transferred in the 80% than 67% criterion cases.

Another possible issue could be that the minimum number of prereversal trials was too low. Because random guessing should result in 50% correct selections, it would not be uncommon for a participant to randomly guess 7 out of the first 10 problems correctly. In the 67% criterion, this would trigger the reversal and the posterior trials. Such a case is a false positive of rule acquisition. Subsequently, the TI for that problem would not be measuring the ability learn, change, or suppress a rule. These false positives would not occur in all cases, yet they may have occurred frequently enough to increase the variability of the results. This could be reduced by increasing the minimum number of prereversal trials. Unfortunately, this might cause some problems to be learned to a greater extent than desired, especially in the 67% case. To prevent this, a more complex problem could be utilized (i.e., one that has 5 digits instead of 4). As the complexity and minimum number of trials are increased, the likelihood of a participant reaching criterion "on accident" would decrease.

Another way to reduce the effects of random successes while maintaining low criterion levels would be to increase the number of choices. If there were three answers from which to choose instead of two, then the chance of choosing correctly would be lowered from 50% to 33%. However, part of the benefit of the TI paradigm is that it measures critical levels of underlearning. As such, it is necessary that the rule not be overly learned. As expected random performance and actual performance diverge, so diminishes the level of under-learning.

Therefore, criterion level of performance should probably be set in a way that reflects the level of significance desired.

It is possible that there were extraneous distractions in the learning reversal task. For example, the binary strings were presented at random vertical locations on the screen. Participants who were more visuo-spatially oriented may have tried to sort by the vertical location instead of by the prototype. With a greater number of variables to test, such participants may have required a greater number of trials to learn the correct rule. Even though rule acquisition may have taken longer, the participant should have been sorting by the prototype when the criterion level was reached. It is uncertain how post-reversal performance would be affected; however, a large effect is unlikely because the participant would quickly eliminate extraneous sorting criteria after the first few successful trials.

This experiment contained unique research that is important for several reasons. First, it examined TI and set switching in normally functioning, young adults. Previously, most of the data available were generated by nonhuman animals and by humans with mental retardation. Because relatively few errors would have been expected from the college-aged participants, the WCST was manipulated in an attempt to increase the number of perseverations. This was accomplished via a computer implementation of the task, which enabled it to be administered in a speeded way via a shortened stimulus onset. New stimuli were presented immediately following the feedback from the previous selection, which itself was brief. The fast-paced format was desirable for testing the college population, in which few perseveration errors would otherwise be expected. The computer-testing platform also allowed for improved experimental control over the previous paper card version.

Even though this experiment yielded unexpected results, they were nonetheless important. Two cognitive mechanisms that were presumed to be quite similar were shown to be significantly different. The ability to modify learning to solve novel problems stands out as the leading candidate for this difference. It could be that our ability to consciously go against the force of habit and instead rely on abstract rules is part of what we prize as intelligence. It could also be this same ability that has given humankind an evolutionary advantage by enabling us to find creative solutions in an ever-changing environment. If this is an important cause of our success, then it would be worthwhile to determine the best way to measure it and teach it as a skill.

References

Aristide, S., Bernardo, P., Grazia, S., and Gaspare, G. (2006). *Fluid Intelligence and Executive Functions: New Perspectives*. Hauppauge, NY: Nova Science Publishers.

Baddeley, A. D. (1993). Working memory or working attention? In A. D. Baddeley & L. Weizkrantz (Eds.), *Attention: Selection, Awareness and Control. A tribute to Donald Broadbent* (pg. 152-170). Oxford: Oxford University Press.

Baddeley, A.D., Hitch, G.J. (1974). Working Memory, In G.A. Bower (Ed.), *The psychology of learning and motivation: advances in research and theory* (Vol. 8, pp. 47-89), New York: Academic Press.

Bunch, M. E., (1941) The measurement of retention by the relearning method. *Psychological Review*, 48, pp 450-456.

BF1.P7

Colom, R., Abad, F., Rebollo, I., & Shih, P. (2005). Memory span and general intelligence: a latent-variable approach. *Intelligence, Vol 33(6),* Nov-Dec 2005, pp. 623-642.1

Dempster, F. N. (1991). Inhibitory processes: A neglected dimension in intelligence. *Intelligence, 15*, 157-173.

Gardner, H. (1983). *Frames of mind: The theory of multiple intelligences*. New York: Basic Books.

Kane, H. & Brand, C. (2007). The variable importance of general intelligence (g) in the cognitive abilities of children and adolescents, *Educational Psychology, Vol 26(6),* Dec 2006, pp. 751-767.

Kolata, S., Light, K., Grossman, H, Hale, G., & Martzel, L. (2007). Selective attention is a primary determinant of the relationship between working memory and general learning ability in outbred mice. *Learning & Memory*, Vol 14(1-6), Jan-Jun 2007, pp. 22-28.

Logie, R. H. (1995). *Visuo-spatial working memory*. East Sussex, UK: Lawrence Erlbaum Associates.

Miller, E.K. & Cohen, J.D. (2001). An integrative theory of prefrontal cortex function. Annual Review Neuroscience. 2001; 24:167-202.

Miyake, A., Friedman, N. P. *et al*. (2001), How are visuospatial working memory, executive functioning, and spatial abilities related? A latent variable analysis, *Journal of experimental psychology: General*, 130(4), p 621-640.

Rajalakshmi, R., Jeeves, M. A., Discrimination learning and reversal learning in children as related to performance on certain WISC items and on the Goodenough Draw-A-Man test. *Journal of genetic psychology*, 106(1), 1965. pp 149-156.

BF699.P43

Rumbaugh, D. M. (1969). The transfer index: an alternative measure of learning set. Proceedings of the 2nd Congress of International Primalological Society 1, 267–276.

Rumbaugh, D. M., & Pate, J. L. (1984). The evolution of cognition in primates: A comparative perspective. In H. L. Roitblat, T. G. Bever, & H. S. Terrace (Eds.), *Animal cognition*. Hillsdale, NJ: Erlbaum.

Spearman, C. (1904). "General intelligence" objectively determined and measured. *American Journal of Psychology, 15,* 201-293.

Sternberg, R. J. (1985). Beyond IQ: *A triarchic theory of human intelligence*. New York: Cambridge University Press.

Towse, J. & Houston-Price, C. (2001). Reflections on the concept of the central executive. *Working memory in perspective*, Andrade, Jackie (Ed); pp. 240-260. New York, NY: Psychology Press.

Figures

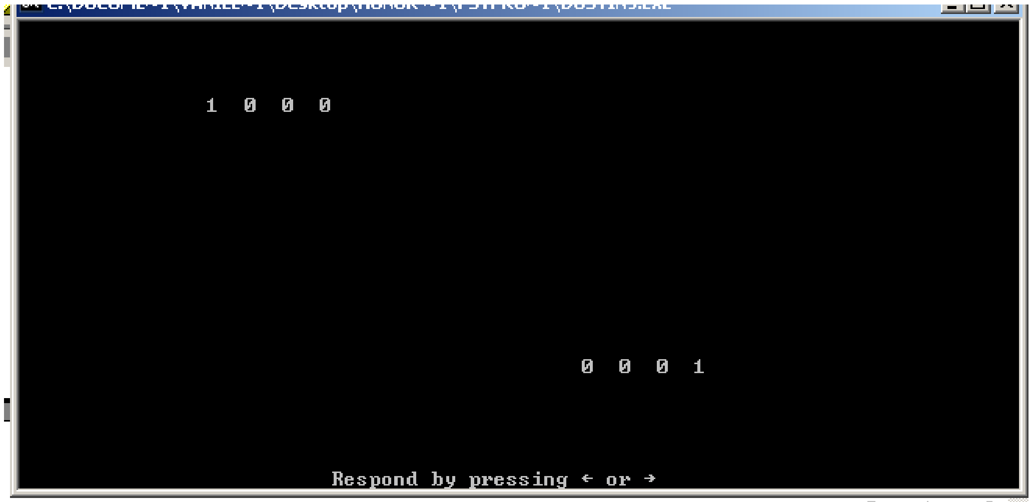

Figure 1. Learning Reversal Task screen shot.

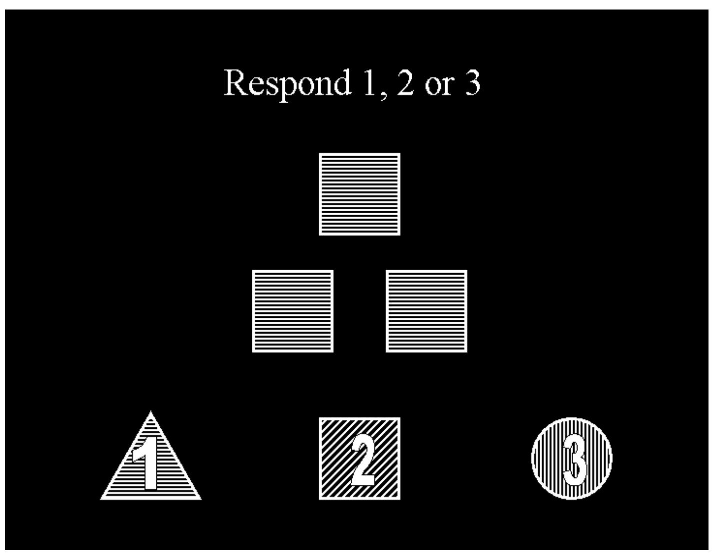

Figure 2. Wisconsin Card Sort Task (WCST) screen shot, with pattern substituted for color.

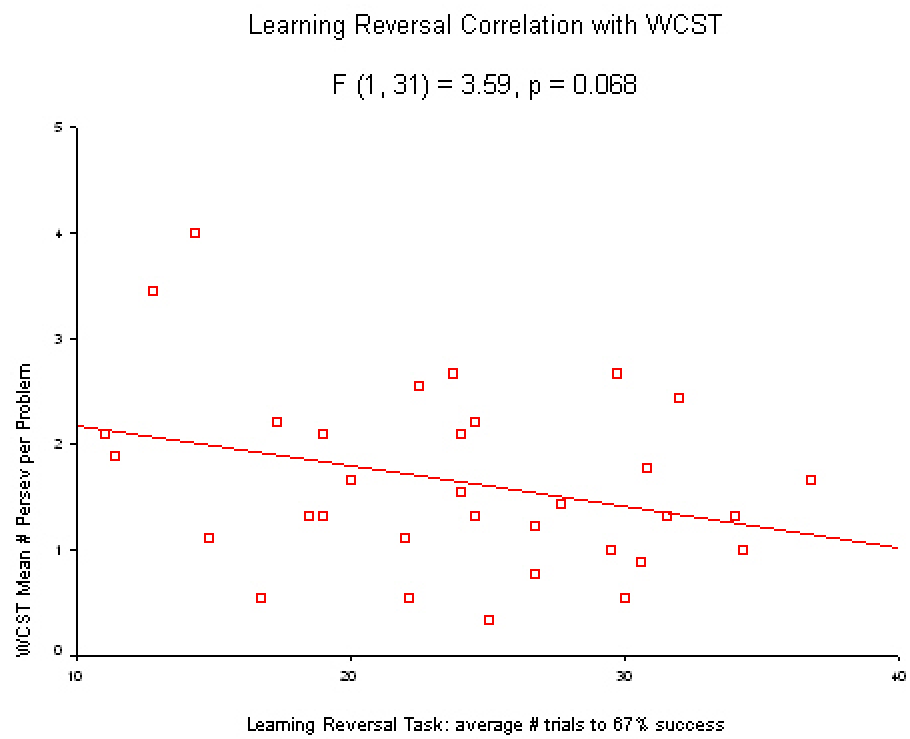

Figure 3. Pre-reversal learning as indicated by the number of trials to reach the 67% criterion level approaches significance in predicting the mean number of perseverations on the WCST

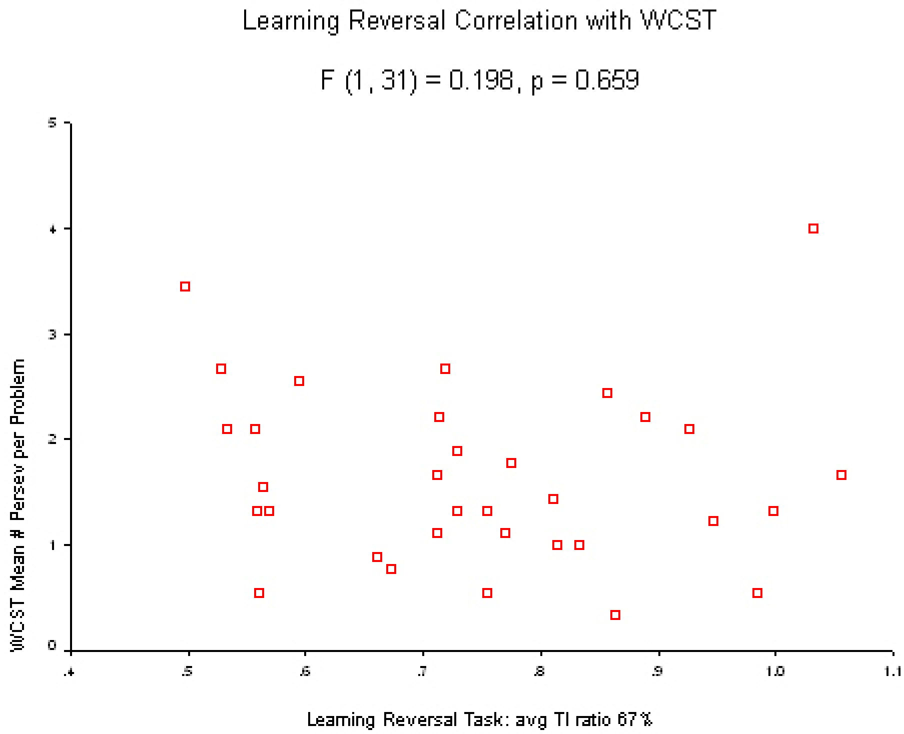

Figure 4. Learning reversal ability as measured by transfer index does not predict executive functioning as measured by the WCST